4

*Local Journalism
without Journalists?
Metric Media
and the Future
of Local News*

ASA ROYAL

PHILIP M. NAPOLI

ABSTRACT

Over the past 15 years, more than 2000 newspapers in the United States have shuttered, leaving the country pockmarked by news deserts. As local newspapers have closed, networks of online local news sites have risen to replace them. We assess the largest network of this kind, Metric Media, by scraping and analyzing the content of its 999 outlets over a 2.5-month period. We find that Metric Media prioritizes the publication of state and national partisan content at the expense of local news, algorithmically generates most of its content, and fails to live up to its mission of democratizing and revitalizing local journalism.

KEYWORDS: Local journalism; digital journalism; business models; media economics; economies of scale; Metric Media.

1. INTRODUCTION

Over the past 15 years, more than 2000 newspapers in the United States have shuttered, leaving the country pockmarked by news deserts (Abernathy 2020). Absent newspaper watchdogs, communities in news deserts have seen rises in local government borrowing costs, increases in government wages (Gao, Lee, and Murphy 2020), decreased competitiveness and engagement in local elections (Schulhofer-Wohl and Garrido 2013; Hayes and Lawless 2017; 2021), and the loss of over 36000 journalism jobs (Abernathy 2020).

As newspapers have lost ground, a number of online local news networks have grown to take their place (Nyhan 2019; Mahone and Napoli 2020). One of the largest of them, Metric Media, has launched over 1000 local news sites across the country and has plans to launch an additional 15000 (Graham 2020).

This influx of local news networks, which comes at a time when the economics of local journalism are growing more challenging, would seem to be a positive development for communities and an indicator of a potentially viable economic model for local journalism. However, journalistic reports have raised significant concerns about how well outlets in these new online news networks serve communities' information needs. Many of these networks are plagued by issues such as a lack of transparency in ownership and funding, a lack of original, local reporting, and a willingness to engage in pay-for-play political influence operations (Alba and Nicas 2020; Bengani 2019). Some, including Metric Media, have been cast as mass producers of low-quality "pink slime" journalism (Bengani 2020; Tarkov 2012). Recent reporting has shown that Metric Media collaborates with and is funded in part by conservative and corporate advocacy groups, whose support the network does not disclose (Bengani 2021).

Considering the critiques and concerns these emerging local news networks have generated, their rapid growth highlights the need for greater empirical scrutiny of their operations. Toward that end, this paper presents a case study of Metric Media via a 2.5-month observation of the home pages of the company's 999 news outlets.[1] We chose to study Metric Media because it is the largest network of local news sites operating in the United States and has attracted a considerable amount of journalistic critique and analysis (Nyhan 2019; Alba and Nicas 2020; Bengani 2020). Its size allowed us to conduct a rigorous quantitative analysis of its story production patterns,

_____

[1] On one company page, Metric Media Claims to operate 1300+ "community news sites" but only lists 1,051 (plus nine publications dealing with "industry news"). On another page, the company only lists 967 local outlets. We operated from the latter list combined with another (see "Identifying Metric Media Outlets" below).

while the fact that it has been a focal point of investigations has let us draw on previous reporting to understand the network's operations.

To ground our case study, we scraped and recorded more than 700000 data points on the daily front page composition of every local news site in the network. We also recorded details on the more than 200000 unique stories shown across the network during the observation period. With these data, we address the following research questions:

1. Are there any notable patterns in the placement of Metric Media sites, content, and journalists across the country? Do they fill gaps left by the demise of print papers?
2. How robust is the journalism produced by Metric Media sites in terms of its timeliness, originality, local orientation, and use of human reporting?
3. Is there evidence in Metric Media's reporting patterns of coordinated political influence efforts?

2. LITERATURE REVIEW

In his field-defining work, *Media Performance: Mass Communication and the Public Interest,* Denis McQuail (1992, 12) describes media performance research as a hybrid of social responsibility and empiricist scholarship, "linking normative principles with a set of research procedures." In the 30 years since the publication of McQuail's book, the need for research that provides a rigorous empirical orientation *and* normative lens to the behavior of our media institutions (particularly our news media) has intensified, particularly within the U.S. context that is the focal point for this analysis.

As the economic and technological environment has evolved, the pressures that undermine a strong public interest orientation for commercial news media in the U.S. have grown more pronounced (Knight Commission on the Information Needs of Communities in a Democracy 2009), as have the pressures to serve a more narrow constituency of news consumers (Usher 2021). At the same time, the technologically-facilitated lower barriers to entry have meant that the long-desired greater diversity of contributors to the news and information ecosystem has, it could be argued, been achieved (Waldman and Working Group on the Information Needs of Communities 2011). However, it could also be argued that this diversity has been achieved at the cost of allowing a far greater array of actors with a questionable commitment to serving the public interest entrée to this ecosystem (Benkler, Faris, and Roberts 2018). These actors have opportunities (thanks to social media platforms) to access unprecedentedly large aggregations of audience attention (Martens et

al. 2018; Rashidian et al. 2019).

Parallel to — and interconnected with — these developments, has been the re-emergence of partisan journalism in the U.S.: first at the national level, and now, increasingly, at the local level. Partisan journalism is hardly a new phenomenon in the U.S., dating back to the earliest days of formalized newspapers (Kuypers 2013). However, as the commercial model of journalism took hold and expanded through the late 19th and into the 20th centuries, overt partisanship diminished as a journalistic norm (Kaplan 2002), leading, ultimately, to a period throughout much of the mid- to late-20th century that was characterized by a relatively strong ethos of non-partisanship throughout most sectors of U.S. journalism — a period scholars have described as the "heyday of traditional objectivity" (Meyers 2010, 143).

And while the very notion of objectivity in journalism has justifiably come under intense re-examination and critique (see, e.g., Wallace 2019), it is difficult to contest that the post-broadcast media age (see Prior 2007) has brought with it a confluence of circumstances that has triggered a return to the more overtly partisan model that characterized the early days of journalism in the U.S. (Abrahamson 2006). This is a process that first took hold at the national level, in contexts such as 24-hour cable news and syndicated talk radio (Jamieson and Cappella 2008; Sherman 2014; Stelter 2020). And now there are indications that, as the local news ecosystem evolves and adjusts to the new economic and political realities in the U.S., hyper-partisanship is becoming more pronounced in local journalism (NewsGuard 2021; Skibinski 2021).

It is within this conceptual and historical framework that we consider the rise of a new digital news network such as Metric Media. As traditional local news sources (particularly local newspapers) have suffered dramatic declines due in large part to their ability to compete for advertising dollars with large digital platforms such as Facebook and Google (Rashidian et al. 2019), the widespread hope has been that new digital news outlets, brandishing innovative revenue models or financial support systems, greater efficiencies in production and distribution, and a greater commitment to engaging with their communities, would arise to take their place, and effectively fill the void left by their predecessors (see, e.g., Carlson and Usher 2016; Jacob 2020; Konieczna 2018). It is too early in this process to tell whether robust and effective replacements for the traditional local newspaper will be able to take hold. However, recent research suggests that digital-only local news sites make quite modest contributions to the totality of original local journalism in their communities and that local newspapers, despite their declining state, continue to provide a disproportionate amount of the original local journalism that serves their communities' information needs (Mahone, Wang, Napoli, Weber, McCollough, 2019).

Indeed, at this point, local journalism research has arguably done a more thorough job of empirically demonstrating the performance declines of traditional local news sources than it has of evaluating how well emergent local news sources are meeting their communities' information needs. This study is an effort to address this imbalance, by bringing the same kind of empirical rigor that has been brought to bear to demonstrate the performance declines of traditional local news sources to the question of how well emergent local news sources are picking up the slack and effectively serving the information needs of their communities.
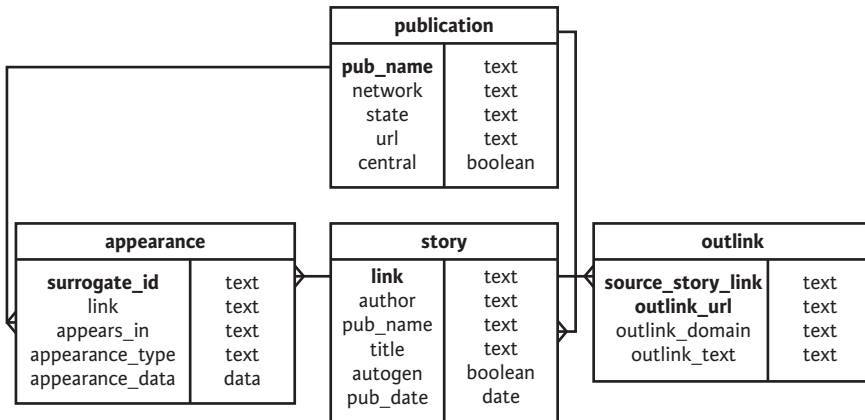
3. METHOD
3.1 IDENTIFYING METRIC MEDIA OUTLETS AND ESTABLISHING A DATABASE SCHEMA

To identify news outlets in the Metric Media network, we relied on one of the company's corporate websites, which lists 965 outlets the company operates (Metric Media n.d.-a). We added to that list 34 Illinois-based news sites, operating as part of a news network named LGIS (Local Government Information Services Network). LGIS sites identically resemble Metric Media sites, and previous research by Priyanjana Begani of Columbia University has found that they use some of the same servers, infrastructure, and logging IDs as the Metric Media sites (Bengani 2019). Further reporting by Bengani has shown that the two networks share funding sources and personnel (Bengani 2021).

We decided to treat the combined networks as one, and with a list of 999 outlets in hand, we manually inspected the sites to determine which data were worth tracking. The resulting data schema (see Figure 1) was intended to keep track of two views of the Metric Media news network: first, a daily record of all stories appearing on each of its outlet's front pages; second, a list of each unique story that appeared across the network.

Figure 1: Database schema

publication	
pub_name	text
network	text
state	text
url	text
central	boolean

appearance	
surrogate_id	text
link	text
appears_in	text
appearance_type	text
appearance_data	data

story	
link	text
author	text
pub_name	text
title	text
autogen	boolean
pub_date	date

outlink	
source_story_link	text
outlink_url	text
outlink_domain	text
outlink_text	text

3.2 COLLECTING DATA

To collect data, we wrote code[2] that visited each site in the Metric Media Network, observed every immediately visible front-page story, and then recorded details in the database described in Figure 1. The scraping code ran every day from November 16, 2020 to February 1, 2021. With our data schema, we were able to conduct longitudinal analyses of coverage across the network, as well as reproduce point-in-time facsimiles of any news outlet's front page on any date during the observation period.

3.3 AUTOGENERATED STORY LABELING

Before we began scraping data, we observed that many stories published by Metric Media outlets looked strikingly similar. These stories, replicated across outlets, often contained identical text, save for locally specific statistics and names that were plugged in based on an outlet's location. Further observation led us to a simple heuristic for identifying such stories: all stories with authorship attributed to "Metric Media News Service" or attributed to a publication by name followed these plug-in templates. We used that heuristic to mark a large subset of stories (~207,000, 95% of the total) as autogenerated[3]. We then developed a set of ~200 regular expressions that matched the titles

2 Code available from authors.
3 We also tagged ~2,500 stories attributed to "Press Release Submission" as auto generated. As the attribution suggests, these pieces were word-for-word press releases reproduced from outside sources.

of nearly all the autogenerated articles[4]. Each regular expressions denoted a group or "strain" of identical autogenerated articles. For example, the regular expression "{} campaign committee refunded ${} in {}", matched a strain of autogenerated articles about campaign committees refunding donations.

Following the data collection period, we discovered roughly 1000 more autogenerated stories by matching story titles in the database against the list of regular expressions. Many of the stories that escaped detection by the author heuristic were missed because their authorship was attributed to a permutation of a publication's name (e.g., a story in *East Central Georgia News* was attributed to the author *EC Georgia News*). Upon detecting those stories, we proceeded to label them as autogenerated.

Finally, during a manual inspection of the data, we noticed two "strains"[5] of stories that were reproduced according to strict templates but were not caught by the author heuristic or title matching (see sections 3.4, 4.4). Those, too, were marked as autogenerated.

3.4 ASSESSING AUTOGENERATED STORY LABELING

To validate our autogenerated story labels, we analyzed[6, 7] intra-strain article similarity. We hypothesized that that given a set of articles we had labeled as belonging to the same autogenerated strain—i.e., articles whose titles matched the same regular expression— the text of each article in the set would closely resemble that of the others.

To test this, we queried our database for a list of up to 200 articles in every strain (or 1% of articles in the strain, whichever figure was greater). For each strain, we randomly picked a "canonical" article from the database sample. Next, we fetched the text of every article in the sample group and compared it to the text of the canonical article using a simple bag-of-words technique: given two articles, we stripped them of proper nouns and stop words[8] and considered each as a set of vocabulary. We then divided the size of the intersection of both articles — the number of words they shared — by the size of their union — their combined unique word count. The resulting metric (Jaccard Similarity) assigned a score of 1.0 to two documents with m and n

---

4 This set of regular expressions matched ~99.96% of the stories we had labeled as autogenerated. Many of the remaining ~.04% (80/211,000) stories we had labeled auto-generated appeared to be human-written articles we falsely classified based on the heuristics described above.

5 Used throughout the paper to refer to a group of autogenerated articles that follow a single template.

6 Code available from authors.

7 18 recorded story strains representing ~17,000 stories appear to have been taken down from Metric websites and were not included in this analysis.

8 As identified by the Python library NLTK

unique words if their vocabularies were identical (*intersection/union = [m+n]/ [m+n]*) and 0 to two articles if they had no words in common (*intersection/ union = 0/[m+n]*).

After running comparisons within each strain sample, we found the average strain had a mean similarity score of 0.92 (median 0.976), a result indicating that most autogenerated stories were basically identical, varying only by a few words.

A discussion of the autogenerated stories follows later in the piece, but we caveat it by admitting that we do not have direct knowledge of how Metric Media's articles were authored or published. While it is true that Metric Media advertises its creation of autogenerated content ("Metric Media" n.d.). and that our analysis found large strains of nearly identical articles in the Metric Media corpus, it is conceivable, even if unlikely, that humans reporters *did* author, for example, the 20,000+ identical stories on local gas prices we observed. It is perhaps more likely that humans had some level of input on the two-story strains we manually labeled as autogenerated (see section 3.4), which totaled ~200 stories. Stories in both strains were attributed to "Staff Reports" — as opposed to the normal autogenerated story bylines — and one strain saw 143 permutations of a template published one day and a single new permutation published four days later, a pattern that seems unlikely to have emanated completely from a script.

3.5 EVALUATING METRIC MEDIA'S COVERAGE OF 2020 ELECTORAL FRAUD CLAIMS

After observing a large amount of content in Metric Media outlets that covered Donald Trump's unfounded allegations of electoral fraud in the United States' 2020 presidential election, we undertook a content analysis to find patterns in that coverage. Specifically, we investigated whether stories about electoral fraud — which Trump alleged had robbed him of victory — were published more often in states whose election results the Trump campaign was challenging, and whether such coverage squeezed out local news in those states during the election "contestation" period. We chose to focus on coverage patterns in the three states won by the narrowest margins in the 2020 election, Georgia, Wisconsin, and Arizona) each of which had its election results challenged by the Trump campaign. In our analysis, we classified every human-written story produced by Metric Media outlets in those states from election day 2020 onwards as covering one of the following topics: electoral fraud; the presidential election; non-presidential elections; or "other", a catch-all topic including anything else (such as local news).

To draw comparisons between Metric Media's content production in the selected trio of contested states and production elsewhere, we randomly

picked ten states outside the trio, conducting an identical content analysis on the stories produced by outlets in those states during the same period.

4. FINDINGS

### 4.1 PLACEMENT/DISTRIBUTION OF METRIC MEDIA OUTLETS

Metric Media news outlets are roughly distributed across states according to population, though a few notable outliers (North Carolina, Ohio, and Pennsylvania) have significantly more outlets than would be expected and one, New York, has significantly fewer (see Figure 2). The three positive outliers were considered potential swing states in the 2020 presidential election (FiveThirtyEight 2020; Wollner 2018).

Figure 2: Metric Media Outlets Serving State by Population
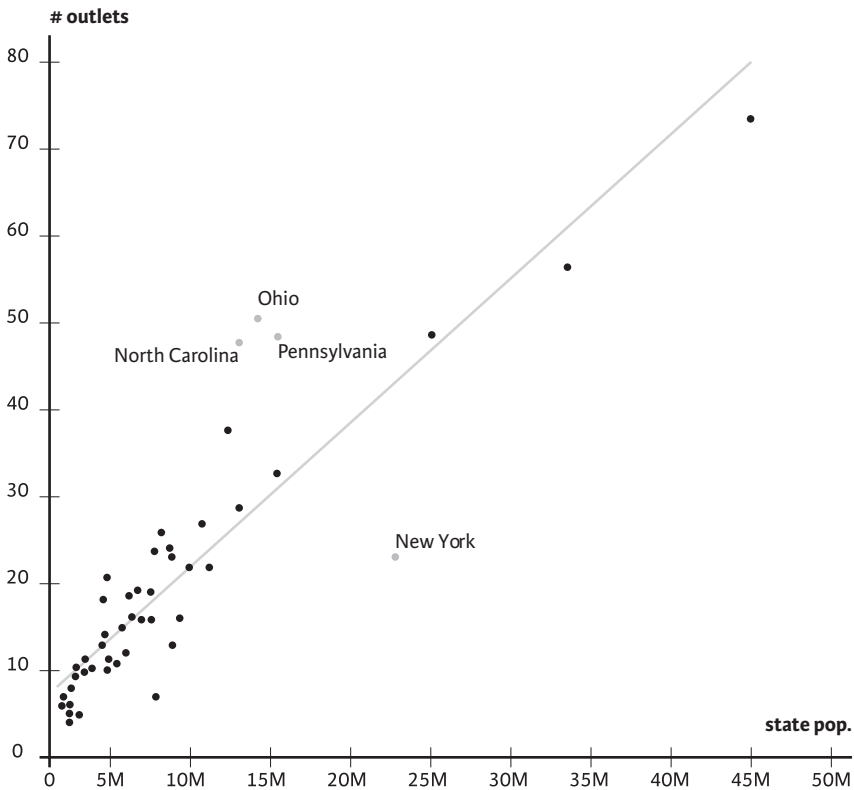
Surprisingly, the number of Metric Media outlets is not closely tied to the total amount of Metric Media's journalistic output in that state (see Figure 3), indicating inequalities in outlets' production.

Figure 3: # of human written stories published in a state v. # of outlets
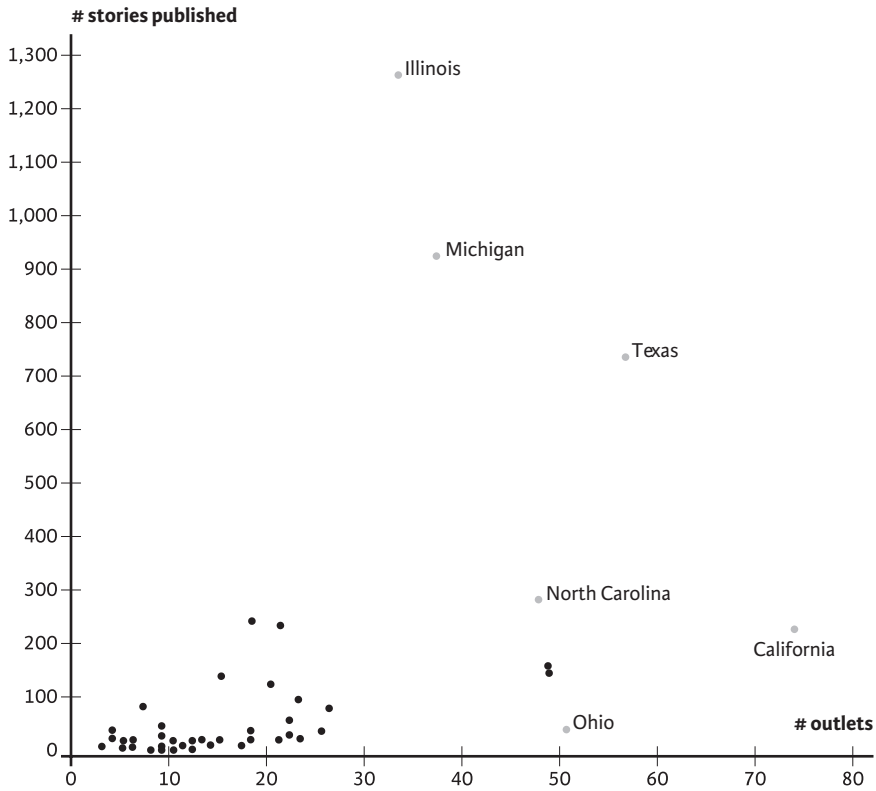
4.2 AUTHORS' OUTPUT AND PLACEMENT ACROSS STATES

During the data collection period, we observed a total of 4,246 non-auto generated (presumably human-written) stories shown across Metric Media's front pages and RSS feeds[9]. These stories spanned 544 publications across 48 states. 455 Metric Media publications never featured a human-bylined article on their front page during the observation period.

The 4,216 non-autogenerated stories were attributed to 161 unique human bylines, but the distribution of authorship volume was highly skewed: the median author published just a single article, while the average author published 26.48 articles. Among the top ten most prodigious authors, the average published 236.5 stories.

_____

9 These included stories with human bylines, as well as stories attributed to "Staff Reports", "{Outlet Name} Reports" etc.

Many authors wrote stories for publications in several states: the 20th percentile author reported for publications across 4 states, the median 9.5, and the 80th percentile 16 different states. Authors' output was even more dispersed at the outlet level (this dispersion serves as a proxy for measuring how many localities individual authors reported in): the 20th percentile author published in 10.4 outlets/localities, the median in 30, and the 80th percentile in 75.6.

4.3 STORY TIMELINESS

Over the course of the data collection period, we found that the median age of front-page stories across Metric Media's 999 outlets was 81 days; with each passing day, that figure rose nearly linearly (see Table 1).

Table 1. Median age of front-page stories across all Metric Media outlets.

DATE	MEDIAN AGE (DAYS)
2020-11-16	40
2020-12-01	54
2020-12-16	69
2021-01-01	85
2021-01-16	100
2021-02-01	110

The aging nature of Metric Media's front-page content was reflected in the number of recently published stories (defined as those ≤ 2 weeks old) present on outlets' front pages. The median publication (in fact, 761/999 publications) had zero "recent" stories, while the average, higher due to outliers (see section 4.5), had 0.49. Only a tenth of publications (97/999) had three or more "recent" stories on their front pages on the last day of our collection period, and less than 5% (42/999) had three or more stories less than a week old. The age distribution of front-page stories across all outlets on our final day of data collection is shown in Figure 4.

Figure 4: Stories on Metric Media front pages grouped by age (2021-02-01)[10]
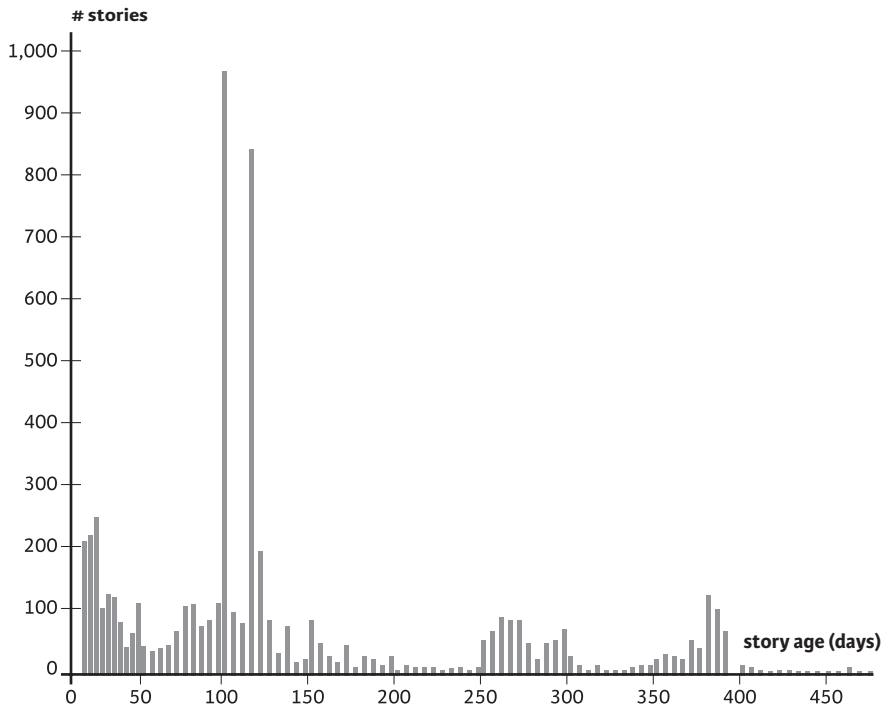
4.4 AUTOGENERATED CONTENT

Metric Media's most prolific authors do not appear to be human. A corporate website notes the company publishes "over 5 million news articles every month, including autogenerated data stories". During our 78-day collection period we recorded the appearance of ~217,000 unique articles, 97% of which we judged to be autogenerated (see section 3.3). Those autogenerated stories, essentially mad-libs scripts, used templates to wrap data pulled from sources including U.S. Census manufacturing reports, Board of Elections donor lists, and local Mass schedules. Some story templates, like a strain reporting local gas prices from GasBuddy, a gas price tracker, were reproduced more than 20,000 times during the collection period and transplanted across nearly all outlets. To give a sense of what autogenerated stories looked like, we have included a selection of titles below:

---

10 The two spikes at ~100 days and ~115 days were caused by the publications of two large auto-generated story strains (see section 4.4).

- "AT&T, Inc. employee Stacey Johnson Batiste contributes a total of $2,000 to Democratic Party candidate Kamala D. Harris"
- "2 professional licenses were issued in Los Angeles during week ending January 16"
- "Amid statewide drop in public school enrollment, Ysleta ISD saw a -5.33% decline"
- "Cory Diez earns $22,188 in compensation from Clark County in 2019"
- "25 deaths from diabetes mellitus in Oklahoma during week ending November 21"
- "Logan CC has an inmate population of at least 1,218 in September 2020"
- "Lancaster Elementary School District had 1,618 seventh grade students during 2018-2019 school year"
- "Paper and paper products wholesalers report 7.6 percent increase in year-over-year November inventories"
- "822 seniors over 79 who were in US test positive for COVID-19 during week ending October 5"

Though most stories (97%) produced by Metric Media outlets were auto-generated, only 1.31% of all autogenerated stories were placed on the front page of any of the network's outlets. The rest either sat in outlets' RSS feeds before being cycled out a few days later or were stowed below the viewport on the front page (see Figures 5 and 6). Even so, about a third of outlets featured exclusively autogenerated stories on their front pages (see figure 7).

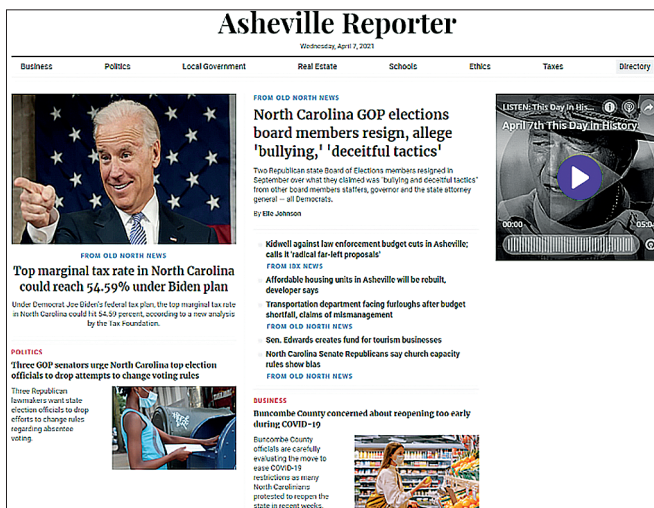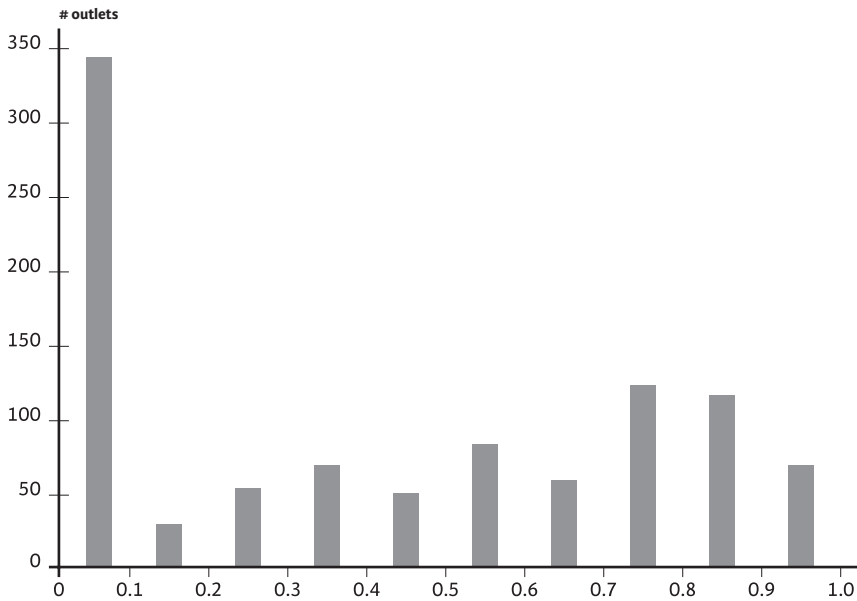Figure 5: Screenshot of the Asheville Reporter front page (immediate view)

Figure 6: Screenshot of the Asheville Reporter front page (scrolled below initial view-
port to show autogenerated stories)

Figure 7: Metric outlets sorted by median fraction of front-page stories that are
human written

Generally, the small share of autogenerated stories that did reach the top of Metric Media outlets' front pages was markedly different than the rest. Instead of featuring tidbits of local information from public datasets, these stories frequently discussed state and national politics. For example, a strain of stories titled "Top marginal tax rate in {state} could reach {x}% under Biden plan" was prominently featured on the front pages of Metric Media publications in all 50 states; another strain, titled "{State} Gov. {Governor name} gets '{A-F}' grade for fiscal management", citing ratings by the conservative CATO institute, was featured on the front pages of outlets across 41 states. Both the Biden tax plan and Cato governor grade stories were published by state hub outlets (see section 4.5). Together, stories in the two autogenerated strains made up just .43% of all autogenerated pieces, but because of how often they were featured across the network, accounted for 64% of all front-page appearances of autogenerated stories.

During the data collection period, we observed the elevation of sevDuring the data collection period, we observed the elevation of several other autogenerated story strains. On November 18, 2020, shortly after Joe Biden's general election victory over Donald Trump, 143 stories ran across Metric's Georgia outlets, highlighting individual counties' absentee ballot rejection rates. The cookie-cutter stories suggested that absentee ballots, the majority of which were believed to be cast in favor of Joe Biden (Doherty et al. 2020), were being rejected at anomalously low rates, a claim echoed by Donald Trump in his quest to discredit Biden's win (Reuters Staff 2020).

Though only ~1% of autogenerated stories were ever featured on the front pages of Metric Media's outlets, the Georgia election pieces had little trouble gaining such traction. A third (47) were featured on Georgia outlets' front pages, and 38 were still present at the end of our collection period, nearly two months after the election.

In a similar episode, following a Republican-led vote in the Ohio Senate to roll back the Ohio Department of Health's quarantine authority, Metric Media's Ohio outlets published 33 templated stories, each detailing how individual state senators had voted on the measure.

## 4.5 OUTSOURCED CONTENT AND HUB OUTLETS

During the data collection period, nearly two-thirds of Metric Media publications (637/999) never published a single human-written article. In place of original journalism, these sites featured articles published by other Metric Media outlets. Such article sharing generally did not take place between geographically neighboring outlets, but rather flowed from a group of "hub" sites to the rest of the network. This collection of hub outlets consisted of

one publication per state, each monikered by a state nickname or descriptor (e.g., Georgia's hub outlet was named *Peach Tree Times*). On an average day, hub outlets were the original publishers of 43% of the stories displayed across every front page in the Metric network despite making up only 5% of the network population. Figure 8 shows the distribution of all 999 outlets by the percentage of their front-page content "borrowed" from hubs[11].

Hub outlets were often outliers to the trend of aging content discussed earlier. Whereas only 11% of all outlets had at least one news story less than two weeks old on the front page, 64% of hub outlets did. Even when the front-page content of hub outlets did age, it usually remained much more up to date than those of other "local" publications in their respective states. We illustrate this in Figures 9 and 10, which graph the median story age across "hub" and "spoke" outlets in two representative states (North Carolina and Pennsylvania).

Though we did not conduct a formal content analysis, a cursory observation of central outlets' front pages makes clear that they deal almost exclusively in state and national political news.

Figure 8: Metric Media outlets sorted by fraction of self-published, human-written content on front page (2021-02-01)
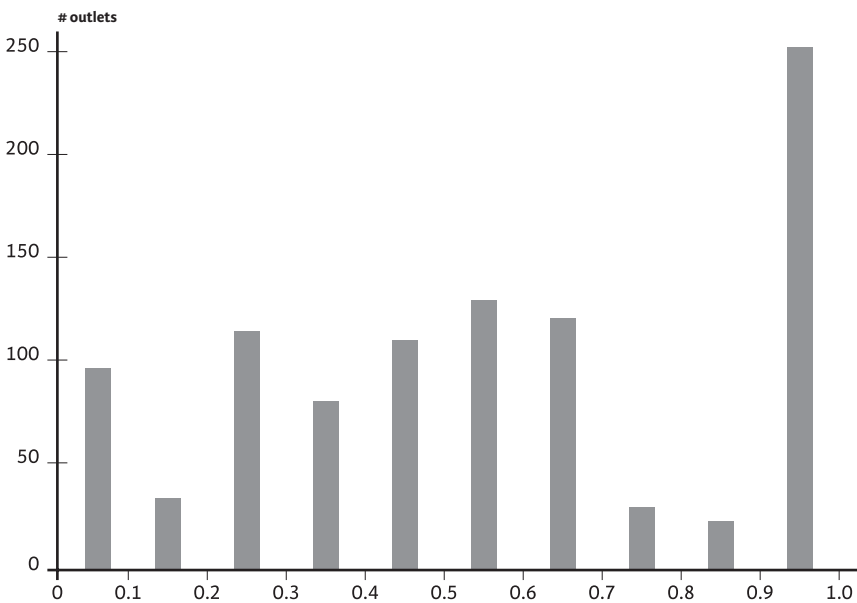
11 Hub outlets themselves make up fifty members of the rightmost bin.

Figure 9: Change in front-page story age over time across NC outlets. Hub outlet highlighted
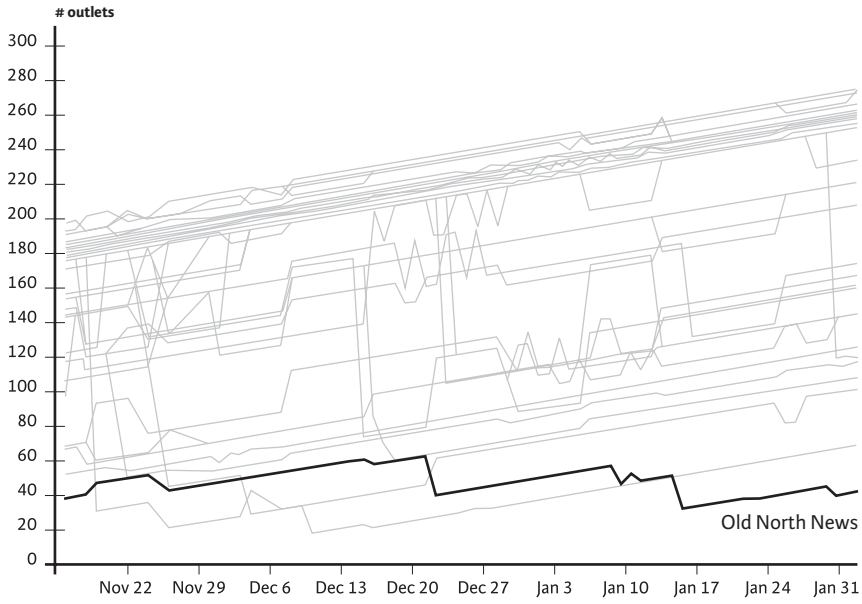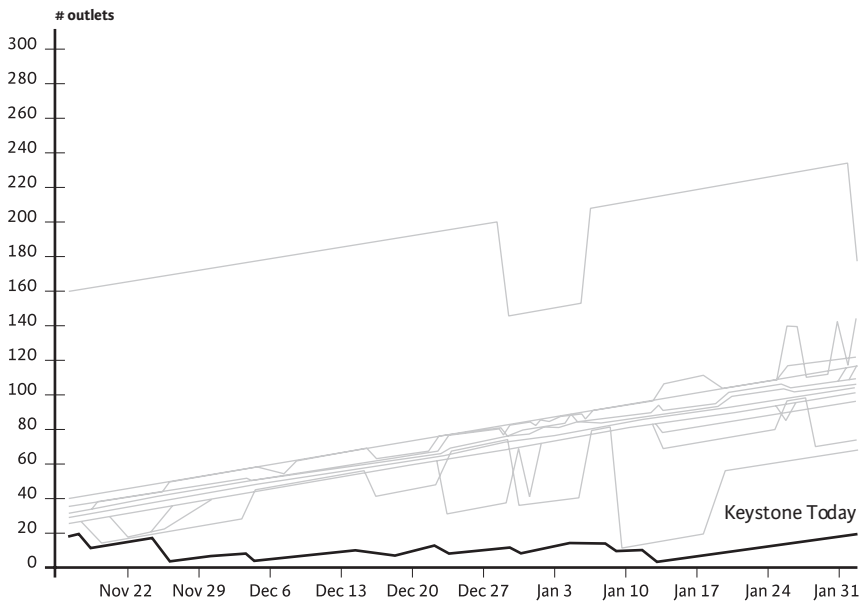

Figure 10: Change in front-page story age over time across PA outlets. Hub outlet highlighted

4.6 METRIC MEDIA OUTLETS' COVERAGE OF ELECTORAL FRAUD CLAIMS

Bengani (2019) and Alba and Nicas (2020) have reported that in the past, Metric Media has given outsized positive coverage to conservative talking points and Republican politicians. We explored this possibility of political influence efforts, using as a case study Metric Media outlets' coverage of Donald Trump's allegations of electoral fraud in the 2020 US presidential Election. Though our observation period began too late track story production before and during the election, it was well fit to observe reverberations after.

The data show that Metric Media news outlets in states with close presidential election margins that the Trump campaign was contesting (Georgia, Wisconsin, and Arizona), began to produce stories about electoral fraud during the election contestation period[12] at a rate much greater than rates seen in outlets from a randomly selected control group of states[13]. Outlets in the trio of "battleground" states also saw much steeper content volume drop-offs after the election contestation period than those from states in the control group, suggesting that overall content production in battleground states had been increased to pump out fraud stories.

During the election contestation period, Metric Media publications in Georgia produced a total of 130[14] human-written stories (94 came from the state's hub outlet, *Peach Tree Times*), of which 102 (78.46%) dealt with electoral fraud. But after the contestation period, from January 9, 2021 to February 1, 2021, the same Georgia outlets produced just 3 human-written stories, reflecting a 94.81% reduction in the number of stories published per day. Similar patterns appeared in Wisconsin and Arizona, in contrast to what we saw in control states (see Table 2).

Table 2. Differences in Contested State Outlets' Content

State	% Content covering electoral fraud (Election to Jan 9)	Stories/day (Election to Jan 9)	Stories/day (Jan 9 to Feb 1)	% Change in stories/day
Georgia	78.46%	2.41	0.13	−94.81%
Wisconsin	71.15%	0.96	0.17	−82.69%
Arizona	76.00%	1.39	0.50	−64.00%
Median control	16.22%	0.69	0.29	−38.59%[15]

12 From election day (November 3) 2020 day until January 8, 2021, a day after Congress officially certified the results of the 2020 Presidential Election.
13 The control group states were Florida, Hawaii, Illinois, Massachusetts, Missouri, Nevada, North Carolina, Pennsylvania, Texas, Virginia, and West Virginia.
14 In addition, as mentioned above, the Georgia outlets produced a series of 143 auto-generated/templated stories highlighting what they described as aberrantly low absentee ballot rejection rates in the 2020 election.
15 The discrepancy between this number and the stories/day figures for the control group is brought on because the control stories/day figures are group medians. Calculating a change

5. DISCUSSION

Our analysis suggests that Metric Media's claim to being "the largest producer of local news content in the nation" (Metric Media n.d.-c) is true only in a narrow sense. The company operates nearly a thousand sites and produces a large quantity of content, but to call most of its content "news" or most of its sites "news outlets" conflates the thin veneer of local coverage Metric Media publishes across the country with an operation dedicated to producing substantive local news.

5.1 METRIC MEDIA'S PLACEMENT OF OUTLETS AND AUTHORS

In a 2012 Poynter Institute article, a reporter summarized a Metric Media executive's comments on the (lack of) need for locally-placed journalists: "young reporters cycle in and out of the areas they cover without knowing much about them. By that logic, why not give that work to someone outside the area ...?" (Tarkov 2012).

Relying on autogenerated stories and distant reporting means journalists will miss important stories, and their reporting may lack local context. As discussed above, the median Metric Media author published stories across 9 states and 30 localities[16]. Given the difficulties of reporting even in a single local context, it is highly unlikely that Metric Media's authors, spread so thin, can produce a genuinely locally oriented news product. Our first research question dealt with the broad distribution of Metric Media's outlets and journalists. In addressing that question, we found that though the company has spread nearly 1000 outlets across the country, including many into news deserts, the bulk of the outlets lie almost bare, dusted with a shimmer of one-off stories, but lacking substantive community-based reporting.

5.2 ROBUSTNESS OF JOURNALISM

In our second research question, we asked how robust Metric Media's journalism was in terms of its age, originality, local focus, and use of human reporting.

5.2.1 Age, Originality, and Local Focus

Metric Media presents its outlets as the online equivalents of daily newspapers. But in all but a select few of them — mostly hub outlets catering to state and national political news — original stories are either nonexistent or stale, a conclusion exemplified by the fact that only one third of the 999

---

from those figures is not meaningful, because it conflates raw production levels across states. The figure we arrived at, -38.59%, was calculated by finding the change in story production within each state, then taking the median of that list.

16 Calculated among Metric Media authors who published at least 5 pieces.

outlets we tracked published a single non-autogenerated story during the 78-day observation period.

It is highly unlikely that a local newspaper or its online equivalent can meet the information needs of its readers if it does not publish for 78 days. And to the point of local focus, for most outlets, as discussed above, there is none. Though state hub outlets often share stories with local Metric outlets, such content lending is not a substitute for original local reporting and does nothing to address local information needs.

A case study in Ohio illustrates the gaps left by Metric Media's sparse coverage. In October 2020, after the company took over the 80-year-old independently owned Mount Vernon News, a water main in Mount Vernon, Ohio burst, leaving an eight-foot crack in the pipe, and spurring a city-wide boil advisory (Pepper 2020). For several days, the Metric Media-owned newspaper failed to cover the water main burst, leaving city leaders scurrying to send out text alerts to constituents. In the words of a Republican city council member, the paper "failed to alert its citizens" and was "basically just a bunch of press releases" (Hendrix 2020).

5.2.2 Autogenerated Content

Most of Metric Media's autogenerated stories are never shown on the front pages of an outlet. Such content — autogenerated and out of sight — like the nearly 7000 pieces we recorded about individuals' professional licenses expiring, made up 97% of all Metric Media stories we observed.

Most of it did not constitute substantive local news. Though some articles had a kernel of interesting data — for example, the percentage of minority students graduating from high schools in a given school district — they lacked any context that might explain the data: they were facts without stories. The low value that Metric Media assesses its own autogenerated content to have is evident in how often the network cites it. Within the text of all human-written Metric Media articles, we observed almost 3600 total links. 99 pointed to stories by Metric Media outlets, but only one pointed to an autogenerated story. That these empty autogenerated stories account for most of Metric Media's content suggests that the network's goals are not fully aligned with the production of robust journalism for local communities.

5.3 ELECTORAL FRAUD COVERAGE

As Donald Trump and his presidential campaign attempted to overturn the results of the 2020 general election on the false pretense of electoral fraud, Metric Media outlets in states the campaign contested gave large amounts of coverage to claims about and battles around electoral fraud. In Georgia,

Wisconsin, and Arizona — the three states carried by the narrowest margins in the presidential election — stories discussing electoral fraud made up an average of 76% of all human-written content from Metric outlets. Stories unrelated to the presidential election made up an average of less than 18% of human-written content in each state.

Beyond taking up coverage space and raising doubts about Metric Media claims to be interested in local news, the company's outsized coverage of electoral fraud revealed a predilection for slanted coverage of increased partisan politics. That finding is not without precedent. A *New York Times* investigation of Metric Media uncovered a pay-to-play political operation in which Republican operatives ordered up articles from "local" Metric sites (Alba and Nicas 2020).

6. CONCLUSION

The news and information ecosystem has evolved in ways that have substantially diminished barriers to entry; have made it easier for news sources to conceal aspects of their ownership and funding sources; and have provided opportunities and incentives for more hyper-partisan and inaccurate reporting. This set of conditions has intensified the need for what we might call *accountability research* in journalism, with researchers serving in a similar watchdog capacity vis-a-vis news organizations that news organizations have traditionally served vis-a-vis government and other powerful institutions.

In applying this accountability lens to Metric Media, our findings suggest that the company operates neither professionally nor economically like a genuine local news network. Most of its outlets do not publish original, human-written content and those that do — primarily state hub outlets — are usually not focused on local news. The freelance journalists Metric Media employs report across several states and localities, which means they are unlikely to accrue the reservoir of knowledge about any locality that enables journalists to become effective watchdogs or providers of community information. For Metric Media, opening another "outlet" in a new community is as simple and cheap as registering a new web domain, but as our data show, that community will often get nothing more than a templated web page sharing autogenerated content and non-local partisan news. An understanding of the shortcomings of Metric Media's approach to local journalism is important in light of both the established size of the network (nearly 1,000 sites nation-wide), its stated plans to launch an additional 15,000 sites, and the rise of similar networks.

In some ways, Metric Media's strategies reflects the efforts of journalism outlets to "do more with less." Metric Media has few journalists, de-emphasizes

in-person reporting (Tarkov, 2012), has turned to low-cost data mining to mass produce stories, and operates under a centralized organizational structure that produces economies of scale (though to a fault—many outlets produce no original news stories). Despite that, it appears incapable of producing even a modest amount of local news in many of its outlets. Its failure to actually "do more with less" speaks to the folly of imagining that local news can currently be sustained through automation, centralization, or removing journalists from journalism.

Meanwhile, Metric Media's overt partisanship highlights the threat untransparent local news networks can pose to democracy. Americans trust local news more than any other journalistic institution (Gottfried and Liedke 2021); by usurping its image, newcomers produce a mirage of trustworthy news behind which, as was the case with Metric Media's coverage of the 2020 electoral fraud, lurks disinformation or pay-for-play content.

Despite the failure (or lacking intent) of Metric Media to live up to its mission of restoring the local news environment, one part of its model hints at a way to help produce substantive local journalism that, while not likely to be taken by the company in question, could be used by others.

Metric Media uses public datasets to fuel the creation of local stories across the nation. Operating with unified resources, individual Metric Media outlets benefit from a constant stream of organized data at almost no marginal cost. It is conceivable that real-world local news outlets might be able to replicate that, joining a consortium or organization set up to mine and parse public data and pipe it to relevant local news organizations.

Right now, if a local outlet cannot train and pay a reporter to parse and mine data, it may not bother gathering it at all, depriving the community of valuable information. A pooled data solution stands to prevent that or, at the very least, free up reporters' time for other stories.

While Metric Media outlets' use of mined data verges on farcical — reporting on the number of salon licenses set to expire in various zip codes — local journalism's use of the data need not be: journalists could easily discern whether a kernel of data was ignorable or worth building a story around.

It is even conceivable that the centralized data pipeline described above might feed data to local news organizations only if the data were either anomalous (e.g., surprisingly low graduation rates at high schools in a city one year) or generally considered newsworthy (e.g., local unemployment rates). Challenges to this approach abound. Data is entropic and messy, the money for a hypothetical data mining organization might be hard to come by, and many newspapers have crossed the Rubicon to closure or irrelevance. But in the absence of a resurgence of genuine local news, it is conceivable that

upstart networks like Metric Media may usurp local news' erstwhile reputa-
tion for honest reporting and use it for ignoble ends.

Much of the recent research on the state of local journalism has focused
on documenting the decline of traditional local news sources and estimating
the effects that these declines are having on various aspects of community
well-being. This study suggests that we should not only be concerned about
the declines in traditional local news sources, but also about some of the new
local news sources that are arising in their place. With future research, we
hope to expand our analytical focus, and to conduct analyses of how left- and
right-leaning outlets compare in terms of their performance serving commu-
nities' critical information needs.

REFERENCES

Abernathy, Penelope. 2020. "Vanishing Newspapers." *The Expanding News Desert* (blog). 2020. https://www.usnewsdeserts.com/reports/news-deserts-and-ghost-newspapers-will-local-news-survive/the-news-landscape-in-2020-transformed-and-diminished/vanishing-newspapers/.

Abrahamson, David. 2006. "The Rise of the New Partisan Press: Forward into the Past." *Journal of Magazine Media* 8 (1). https://doi.org/10.1353/jmm.2006.0007.

Alba, Davey, and Jack Nicas. 2020. "As Local News Dies, a Pay-for-Play Network Rises in Its Place." *The New York Times*, October 18, 2020, sec. Technology. https://www.nytimes.com/2020/10/18/technology/timpone-local-news-metric-media.html.

Bengani, Priyanjana. 2019. "Hundreds of 'Pink Slime' Local News Outlets Are Distributing Algorithmic Stories and Conservative Talking Points." Columbia Journalism Review. December 18, 2019. https://www.cjr.org/tow_center_reports/hundreds-of-pink-slime-local-news-outlets-are-distributing-algorithmic-stories-conservative-talking-points.php/.

———. 2020. "As Election Looms, a Network of Mysterious 'Pink Slime' Local News Outlets Nearly Triples in Size." Columbia Journalism Review. August 4, 2020. https://www.cjr.org/analysis/as-election-looms-a-network-of-mysterious-pink-slime-local-news-outlets-nearly-triples-in-size.php.

———. 2021. "The Metric Media Network Runs More than 1,200 Local News Sites. Here Are Some of the Non-Profits Funding Them." *Columbia Journalism Review*, October 14, 2021. https://www.cjr.org/tow_center_reports/metric-media-lobbyists-funding.php/.

Benkler, Yochai, Robert Faris, and Hal Roberts. 2018. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics.* New York: Oxford University Press.

Carlson, Matt, and Nikki Usher. 2016. "News Startups as Agents of Innovation." *Digital Journalism* 4 (5): 563–81. https://doi.org/10.1080/21670811.2015.1076344.

Doherty, Carroll, Jocelyn Kiley, Nida Asheer, and Calvin Jordan. 2020. "Mail-in Voting, Absentee Ballots and Voter Engagement." *Pew Research Center – U.S. Politics & Policy* (blog). October 9, 2020. https://www.pewresearch.org/politics/2020/10/09/voter-engagement-and-interest-voting-by-mail-and-in-person/.

FiveThirtyEight. 2020. "Ohio President: General Election Polls | FiveThirtyEight." May 14, 2020. https://web.archive.org/web/20200514221932/https://projects.fivethirtyeight.com/polls/president-general/ohio/.

Gao, Pengjie, Chang Lee, and Dermot Murphy. 2020. "Financing Dies in

Darkness? The Impact of Newspaper Closures on Public Finance." *Journal of Financial Economics* 135 (2): 445–67. https://doi.org/10.1016/j.jfineco.2019.06.003.

Gottfried, Jeffrey, and Jacob Liedke. 2021. "Partisan Divides in Media Trust Widen, Driven by a Decline among Republicans." *Pew Research Center* (blog). August 30, 2021. https://www.pewresearch.org/fact-tank/2021/08/30/partisan-divides-in-media-trust-widen-driven-by-a-decline-among-republicans/.

Graham, Jennifer. 2020. "Understanding 'Pink Slime Journalism' and What It Reveals about Conservatives and Liberals." Deseret News. September 7, 2020. https://www.deseret.com/indepth/2020/9/7/21409053/understanding-pink-slime-journalism-and-what-it-reveals-about-conservatives-and-liberals.

Hayes, Danny, and Jennifer Lawless. 2021. *News Hole: The Demise of Local Journalism and Political Engagement.* New York: Cambridge University Press. https://www.amazon.com/News-Hole-Journalism-Engagement-Communication/dp/1108819842/ref=sr_1_1?crid=1M0U5CWF53KBG&dchild=1&keywords=news+hole&qid=1635706289&qsid=140-0122253-5784446&s=books&sprefix=news+hole%2Caps%2C167&sr=1-1&sres=1108819842%2CB09B3WHZ8Q%2CB00133YTM8%2C1736703307%2C0718024133%2C0062976583%2C1790356229%2C0979698847%2C052910069X%2C9004321861%2C0785221085%2C1538136775%2C0785225544%2C0893573019%2C0893573655%2C179712207X.

Hayes, Danny, and Jennifer L. Lawless. 2017. "The Decline of Local News and Its Effects: New Evidence from Longitudinal Data." *The Journal of Politics* 80 (1): 332–36. https://doi.org/10.1086/694105.

Jacob, Mark. 2020. "Are Digital Startups the Great Local News Hope or a Disappointment?" Pew Research Center. https://www.poynter.org/locally/2020/are-digital-startups-the-great-local-news-hope-or-a-disappointment/.

Jamieson, Kathleen Hall, and Joseph N. Cappella. 2008. *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment.* Oxford University Press.

Kaplan, Richard L. 2002. *Politics and the American Press: The Rise of Objectivity, 1865-1920.* Cambridge University Press.

Knight Commission on the Information Needs of Communities in a Democracy. 2009. *Informing Communities: Sustaining Democracy in the Digital Age*. Washington, D.C.: Aspen Institute.

Konieczna, Magda. 2018. *Journalism Without Profit: Making News When the Market Fails.* Oxford University Press.

Kuypers, Jim A. 2013. *Partisan Journalism: A History of Media Bias in the United States.* Rowman & Littlefield.

Mahone, Jessica, and Philip Napoli. 2020. "Hundreds of Hyperpartisan Sites Are Masquerading as Local News. This Map Shows If There's One near You." *Nieman Lab* (blog). July 13, 2020. https://www.niemanlab.org/2020/07/hundreds-of-hyperpartisan-sites-are-masquerading-as-local-news-this-map-shows-if-theres-one-near-you/.

Martens, Bertin, Luis Aguiar, Estrella Gomez-Herrera, and Frank Mueller-Langer. 2018. "The Digital Transformation of News Media and the Rise of Disinformation and Fake News." SSRN Scholarly Paper ID 3164170. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3164170.

McQuail, Denis. 1992. *Media Performance: Mass Communication and the Public Interest.* Thousand Oaks, CA: Sage. https://us.sagepub.com/en-us/nam/media-performance/book202974.

Metric Media. n.d. "Metric Media." Metric Media News. Accessed April 6, 2021a. https://web.archive.org/web/20201101012102/https://metricmedianews.com/.

"————." n.d. Accessed April 7, 2021b. https://web.archive.org/web/20210407191207/https://www.metricmedia.org/.

Meyers, Christopher. 2010. In *Journalism Ethics: A Philosophical Approach.* Oxford University Press.

NewsGuard. 2021. "Special Report: California News Trust Report." NewsGuard. https://www.newsguardtech.com/special-reports/special-report-california-news-trust-report/.

Nyhan, Brendan. 2019. "Americans Trust Local News. That Belief Is Being Exploited." *The New York Times*, October 31, 2019, sec. The Upshot. https://www.nytimes.com/2019/10/31/upshot/fake-local-news.html.

Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections.* 1st edition. New York: Cambridge University Press.

Rashidian, Nushin, Peter D. Brown, Elizabeth Hansen, Emily J. Bell, and Jonathan R. Albright. 2019. "Friend and Foe: The Platform Press at the Heart of Journalism." https://doi.org/10.7916/d8-15pq-x415.

Reuters Staff. 2020. "Fact Check: Georgia Rejected Ballots Did Not Go from 4% to 'Almost Zero' in 2020." *Reuters*, November 23, 2020. https://www.reuters.com/article/uk-factcheck-georgia-rejected-ballots-fo-idUSKBN2832CM.

Schulhofer-Wohl, Sam, and Miguel Garrido. 2013. "Do Newspapers Matter? Short-Run and Long-Run Evidence From the Closure of The Cincinnati Post." *Journal of Media Economics* 26 (2): 60–81. https://doi.org/10.108

0/08997764.2013.785553.

Sherman, Gabriel. 2014. *The Loudest Voice in the Room: How the Brilliant, Bombastic Roger Ailes Built Fox News—and Divided a Country.* Random House.

Skibinski, Matt. 2021. "Pennsylvania Media Trust Report Examines Trust and Integrity for News Coverage in Pennsylvania." https://www.lenfestinstitute.org/local-journalism/pennsylvania-media-trust-report-examines-trust-and-integrity-for-news-coverage-in-pennsylvania/.

Stelter, Brian. 2020. *Hoax: Donald Trump, Fox News, and the Dangerous Distortion of Truth.* Simon and Schuster.

Tarkov, Anna. 2012. "Journatic Worker Takes 'This American Life' inside Outsourced Journalism | Poynter." Poynter. June 30, 2012. https://web.archive.org/web/20140702081758/http://www.poynter.org/latest-news/top-stories/179555/journatic-staffer-takes-this-american-life-inside-outsourced-journalism/.

Usher, Nikki. 2021. *News for the Rich, White, and Blue: How Place and Power Distort American Journalism.* Columbia University Press.

Waldman, Steven, and Working Group on the Information Needs of Communities. 2011. *The Information Needs of Communities: The Changing Media Landscape in a Broadband Age.* Washington, D.C.: Federal Communications Commission.

Wallace, Lewis Raven. 2019. *The View from Somewhere: Undoing the Myth of Journalistic Objectivity.* University of Chicago Press.

Wollner, Adam. 2018. "'The New Swing States:' Presidential Battleground Map Shifts Heading into 2020." McClatchy Washington Bureau. December 12, 2018. https://www.mcclatchydc.com/news/politics-government/election/campaigns/article222917745.html.

PHILIP M. NAPOLI
is the James R. Shepley Professor of Public Policy in the Sanford School of Public Policy at Duke University, where he is also the Director of the DeWitt Wallace Center for Media & Democracy and the Senior Associate Dean for Faculty and Research. His most recent book is Social Media and the Public Interest: Media Regulation in the Disinformation Age (Columbia University Press, 2019).

ORCID www.orcid.org/0000-0001-7308-3044

ASA ROYAL
is an Associate-in-research at Duke University's Sanford School of Public Policy. He is broadly interested in finding new lenses and methodologies to explore the impact of technology on democratic norms and institutions.

Book reviews

BY JOANA LOPES